Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9
SECRET

More on the theory and practice of estimating.

ON THE ACCURACY OF NATIONAL INTELLIGENCE ESTIMATES

Abbot E. Smith

Whenever I talk about National Intelligence Estimates to an intelligence training course, or to any other group, someone always asks: How accurate have these estimates been; what is your score? The question is perfectly legitimate but my answer is usually vague and unconvincing. The purpose of this article is to try to explain why the answer is so unsatisfactory, and then to explore the problem further.

It would seem reasonable to suppose that one could get a truly objective, statistical verdict on the accuracy of estimates. Go through the papers, tick off the right judgments and the wrong ones, and figure the batting average. I once thought that this could be done, and I tried it, and it proved to be impossible. The reasons are various.

The Number of Estimates

Since National Intelligence Estimates began to be produced by their present methods in late 1950, there have been some twelve or fifteen hundred of them. Each of these papers, however, contains a multitude of "estimates," that is, of statements setting forth an explicit or clearly implied judgment. Many of them also include one or more footnotes of dissent, conveying an opinion in conflict with the judgment in the text. I am sure that if one were to try and work out an accuracy score covering the product of nearly twenty years he would have to scan not less than 25,000 judgments, and probably far more. Even if one tested no more than ten or a dozen NIE's he would find several hundred statements to be checked.

Most of these are restricted judgments, frequently appearing in subordinate clauses, and usually introduced because they contribute background to a more contentious or consequential estimate. Most of them were probably not questioned or discussed. I would guess that the vast preponderance of them were quite correct. And if we assume for the moment that they could all be checked, and that 95 percent of them did in fact turn out to be right, I still doubt

MORI 603841
(RIFPUB)

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

that we would be justified in swelling with pride. Most of them were simply too easy. Although indubitably matters of judgment, they were not matters of difficult judgment. In short the batting average, if it were arrived at, would be worth about as much as the batting average of a major league team playing against a scrub outfit in a sandlot. This is why a complete, objective, and statistical tally would not be worth doing.

To be sure, we must not presume that because an estimative judgment appears in a subordinate clause it is necessarily inconsequential. Consider a sentence beginning: "Since the Soviet leaders will not in the near future cease to distrust the United States . . ." If this clause should prove wrong, not only would the rest of the sentence be unsound but the foundations of most estimates about Soviet policy would be undermined. Nevertheless, this is not a judgment which anyone would score high on a list of estimative triumphs. But suppose again (as might well have happened) that sometime in 1958 or so a sentence had begun: "Since no change is to be expected in the Sino-Soviet relationship . . ." Such a clause would certainly in hindsight rank high on a list of egregious errors, yet it is not likely that in 1958 it would have been seriously questioned.

Common sense tells us that a box score of estimates must be selective if it is to mean much; it must take account only of the *important* judgments. In saying this, however, we have left behind the wholly objective approach. Doubtless there are many estimates which everyone would agree to be important, but there are many others on which opinions would differ. The hard fact of life is that the high-level consumer of NIE's—the only person whose opinion really matters—is apt to judge the whole output on the basis of two or three estimates which strike home to him. If they prove correct, NIE's are good; if incorrect, they are bad.[1]

[1] Incidentally, from a strictly professional point of view the intelligence estimator would often rank his successes and failures differently from the way the consumer would. For example, I know of several difficult estimates which proved wrong, and wrong because they showed a failure to grasp the nature of forces at work in a situation; these grieve me greatly, though so far as I am aware no high-level consumer ever noticed them. And there have been some which received high praise, but gave me little satisfaction; they were too easy, or they were merely lucky.

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

The Difficulty of Checking

A great number of the judgments rendered in NIE's cannot be checked at all as to validity; the facts are not available. This is bound to be so; it is no reproach to intelligence collection or research. We estimate, for example, that political leader X is in serious trouble, but then it turns out that nothing much comes of it, and we may never know whether he really was in trouble, or, if he was, whether it was serious. Or we estimate that if the United States undertakes a given course of action the response of other countries will be such and such; but the United States never undertakes that action, and we never know whether we were right or wrong. There are of course a great number of "contingency" estimates, in sentences beginning: "If such and such happens, then so and so will probably follow." But the contingency never occurs, and the estimate can never be objectively checked.

Often those judgments which can be checked have to be scored as partly right and partly wrong; we would view them as "right on the whole," or as "wrong by and large." Or again, suppose we have made an imprudently precise estimate, as that the Soviets will at a given time have 500 missiles of type X, and then they turn out to have 510. Conceivably this might be an important error; more likely it would be considered negligible. But how many more than 500 would they have to have before the estimate should forthrightly be deemed wrong?

Estimative Formulations

The drafters of estimates are deeply conscious of two obligations: to distinguish between statements of estimate and statements of fact, and to convey as clearly as possible the degree of confidence with which an estimate is delivered. On the second point Sherman Kent has written in this periodical. His injunctions may be simplified as follows: since the degree of confidence must usually be conveyed in words, these words should as far as possible be uniformly used and with full understanding of their meaning; for example:

a. Something "is possible" or "may be" true. This constitutes no judgment of probability; it is in effect a statement merely that the thing under consideration is not out of the question. But the fact that it is mentioned at all constitutes a judgment that it is something worth bearing in mind.

b. Something is "probable" or "likely"; this means that there is about a 60 or 65 percent probability of it's occurring or being true.

Let us see how this affects the matter of scoring.

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

First, suppose that an NIE says that "it is possible" that such and such may occur, and then it occurs. We could score this as a correct estimate, which it was. But since a very large number of things are "possible," was it really the kind of judgment that deserves to register a plus for the perspicacity of the estimators? Perhaps it was, and perhaps it wasn't; that will depend on what we were talking about.

Now suppose that the NIE says that something will "probably" occur, and it does not. The estimate was strictly not 100 percent wrong, for it only gave the event about a 60 percent chance of occurring; perhaps it should be scored as 60 percent wrong. But pause a moment, and suppose that somehow we come to realize that there never had been any appreciable chance of the event occurring; then the estimate was really about 100 percent wrong. Or suppose that we come to know that there was indeed a 60 percent chance of its occurring but that something happened—perhaps even an act of US policy taken as a consequence of the NIE—which prevented it from occurring; then the estimate was 100 percent right—or was it?

It ought to be observed that while the subtleties of the preceding paragraph complicate the problem of making an objective and statistical study of the validity of NIE's they are of no consequence in real life. The high-level consumer pays little heed to qualifications. If he is interested in a judgment that something "probably" will happen, and if it turns out not to happen, he denounces the estimate as 100 percent wrong, period. The saddest example of this was seen in the ill-starred estimate of 19 September 1962, issued as the Cuban missile crisis approached. That paper discussed at some length the possibility that the Soviets would put "offensive" surface-to-surface missiles in Cuba. Nowhere does the estimate declare even that the Soviets would "probably" not do so; the presentation was obviously labored, difficult, and inconclusive. Yet the late Senator Robert Kennedy, after the dust had cleared away, wrote as follows in his book, *Thirteen Days:*

> "No one had expected or anticipated that the Russians would deploy surface-to-surface ballistic missiles in Cuba.
>
> "No official within the government had ever suggested to President Kennedy that the Russian build-up in Cuba would include strategic missiles. . .
>
> "The last estimate before our meeting of the 16th of October was dated the 19th of September, and it advised the President that without reservation the U.S. Intelligence Board, after consideration and examination, had concluded that the Soviet Union would not make Cuba a strategic base. . ."

This brings me to the next point.

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

The Discrete Statement and the Context

Neither Senator Kennedy nor the many others who condemned that NIE on Soviet missiles in Cuba were altogether wrong in doing so. The text of that paper was labored and inexplicit. I think that a reader might well have understood that it showed the intelligence community to be beset by the gravest doubts and concerns. Nevertheless it conveyed an unmistakeable impression that the Soviets would probably not do what they did. One may well say that in drafting those passages we ought to have followed Sherman Kent's edicts and come out with a clear-cut statement that the act was improbable; as it turned out we might as well have been killed for a sheep as a lamb. But it was the weight and impact of the context that carried the judgment, rather than any explicit statement. What the estimators probably wanted to convey was something like this: "We really think it unlikely that the Soviets will do this thing, because it would be out of accord with their conduct of affairs in the past, and probably turn out to be disastrous for them; nevertheless, with the evidence as it is, and bearing in mind the gravity of the matter, we think that the risk of their doing it is so great that the US Government should provide for the contingency that it may happen." My concern at the moment is with the question: Supposing that the estimate had in fact said these words or their equivalent, how would its validity have been objectively scored? Still, I suppose, as wrong.

Most NIE's are not so dramatic in their implications, yet a great many convey their message by the context, or rather by the total text. They are something more than collections of discrete statements. Many address questions such as these: what is the situation and what the prospects in country X; what is the trend of Soviet military policy; what is the nature and dimension of revolutionary potential in Latin America; and so on. The validity of such papers depends only partly upon the accuracy of each particular statement in them. It must also be judged by the impact and tone of the document as a whole—the choice of facts which are cited, the distribution of emphasis, the cogency of argument, even the literary quality. I think that such a paper could be basically correct even though it had a great many statements which proved incorrect, and basically wrong even though many statements were accurate.

Sophisticated estimating indeed ought almost always to be something more than bald prediction. The course of events is seldom

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

inevitable or foreordained, even though hindsight often makes it look that way. A good paper on a complicated subject should describe the trends and forces at work, identify the contingent factors or variables which might affect developments, and present a few alternative possibilities for the future, usually with some judgment as to the relative likelihood of one or another outcome. Occasionally such a paper can afterwards be deemed precisely "accurate"; more often it will be difficult to arrive at a verdict in any fashion which can in the strictest sense be called objective. It may be a very long time indeed before we "know" the causes and background of great events. We still get a new analysis, every year or so, of the forces that led to the American Revolution; how soon shall we arrive at objective truth about the forces currently at work in Southeast Asia?

What it comes to is this: a complete, objective, statistical audit of the validity of NIE's is impossible, and even if it were possible it would provide no just verdict on how "good" these papers have been. Like the Bible, the *corpus* of estimates is voluminous and uneven in quality, and almost any proposition can be defended by citations from it. Obviously, if we are to make estimates at all we shall sometimes make wrong ones. An assiduous and hostile critic could certainly make up an extensive list of errors, some of which would be grievous. And a friendly compiler could counter with a massive collection of correct judgments. I usually say to the training course that, being knowledgeable about the contents of NIE's, I believe that on the whole they have been "good." But it may well be thought that mine is a biased verdict, and moreover that since I am a maker and not a consumer of estimates my opinion does not matter anyway.

Seldom if ever does a consumer of consequence pronounce on the virtue of NIE's as a whole, though comments on particular papers or particular judgments have been frequent. The more emphatic of these comments are almost always adverse, since attention seems more likely to be gripped by an important estimate that has gone sour than by one that has turned out right. This is natural enough; it distresses but does not astonish the estimator. Once in a while, however, the temptation to some sort of rejoinder is almost irresistible, and in the following section I indulge myself.

On 1 August 1969, Senator Thomas J. Dodd delivered a speech in the Senate during the debate on the Safeguard program. A part of this

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

speech was devoted to the achievements, or non-achievements, of US intelligence, and the theme was essentially in the following sentence:

> The American intelligence community, although it has performed well in certain situations, has not been impressive when estimating the intentions and plans of our adversaries.

The Senator went on to support this contention by a list of specifics, beginning with the failure to warn of the North Korean Communist attack on South Korea, the subsequent intervention of the Chinese Communists, and the earlier Soviet initial explosion of the A-bomb. Leaving these aside (because they occurred prior to the existence of the present machinery for coordinating National Intelligence Estimates) let us examine some of the others.

a. The intelligence community "failed to predict . . . accurately the Soviet H-bomb."

> Our performance in this respect represents in fact one of those many instances where we were either good or bad, depending on the way one looks at it. We did fail to predict it "accurately." Yet an estimate in March 1953 said that field testing of a thermonuclear device was possible by mid-1955, and further that it would be unsafe to assume that the Soviets would not have a workable thermonuclear weapon by mid-1955. On 18 August 1953, another NIE said that field testing might occur at any time. Soon afterward it was confirmed by analysis that the first test had in fact taken place on 12 August.

b. "In 1956 [the intelligence community] failed to alert us to the Soviet invasion of Hungary . . . And, despite warning signs which many of our lay experts took seriously [it] was also disposed to discount the possibility that the Red Army would invade Czechoslovakia to depose the Dubcek regime."

> It is true that neither the invasion of Hungary in 1956 nor that of Czechoslovakia in 1968 were forecast in National Intelligence Estimates, which represent the consensus of the intelligence community; in fact no such coordinated papers were prepared on these situations in the months immediately preceding the invasions. In both cases, however, and especially that of Czechoslovakia, various estimative memoranda and current intelligence publications reported the state of high tension and the Soviet military build-up. Without saying that invasions were likely, these papers emphasized that they were

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

possible, and were surely under consideration by the Soviet leadership. The US Government was made aware that the invasions might occur, though it was not assured that they would occur.

c. "In the period immediately before the Cuban missile crisis, the advance consensus of the intelligence professionals was that the Soviets would not tempt the fates by deploying nuclear missiles in Cuba."

I have discussed this above, concluding that despite various qualifications that might be made, the Senator's verdict is essentially correct. With respect to the performance of the intelligence community, however, an additional quotation from Senator Kennedy's book is appropriate: "The important fact, of course, is that the missiles were uncovered and the information was made available to the government and the people before missiles became operative and in time for the US to act."

d. "In 1957, the intelligence community was completely without advance information on the Soviet Sputnik."

Strictly construed, the Senator's words seem to condemn the results of collection rather than of estimates, and in this sense they may be correct. Nevertheless, in December 1955 an NIE said that the Soviets could put an earth satellite into orbit by 1958, and in March 1957 another NIE estimated that they could do so by the end of the year. They did, in October. We have always considered this a praise-worthy example of good estimating on the basis of very scanty informatoin.

e. "[After 1957] our intelligence community lapsed into one of its very rare periods of overstatement when it advised the Eisenhower administration that there was a massive missile gap between the Soviet Union and ourselves. Today it has been documented that the so-called missile gap was a Soviet-engineered hoax, and that our intelligence community fell for phony information put out by Khrushchev for the purpose of intimidating us."

We certainly overestimated the number of Soviet ICBM's which would be operational around 1961. But we certainly did not fall for a phony plant by Khrushchev. There was virtually no hard information available, beyond the fact that the Soviets

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

had successfully tested an ICBM in 1957. The principal basis for the overestimate was probably the opinion of the best US missile experts in those early days as to the number of ICBM's that *could* be manufactured in a given period of time, granting a previous successful test. Nevertheless, the estimates were wrong.

f. "In more recent years, conversely, . . . estimates of Soviet intentions regarding the size of Soviet ICBM forces have turned out to be woefully conservative."

A just criticism, despite a few defenses that could be put up.

These exhaust the Senator's list of specifics. Consider now some further general observations which he made: the following quotation combines three passages which were separate in his speech:

> "When it comes to estimates of Soviet intentions, however, there is admittedly a lot of guesswork involved. . . I think it pertinent to point out in this connection that our intelligence community has erred far more frequently on the conservative side than otherwise in their estimates of Soviet capabilities and intentions. . . over and over again, the Soviet performance in the field of armaments has either surprised us completely or substantially surpassed our estimates."

As I have tried to show in preceding parts of this article, it would be idle to attempt to prove or disprove these statements by objective and statistical analysis. With respect to numbers of Soviet weapons, one could easily make up a list of projections which were too low, another of those which were too high, another of those which were substantially correct, and a final one—very short—of those which, thanks more to luck than wisdom, were precisely correct. The projection of numbers, however, is the most precarious of all estimative exercises; there is indeed "a lot of guesswork involved," especially as one looks beyond the two or three years subsequent to the date when the estimate is written.

Suppose we try one test, however, using the somewhat non-objective criterion of "importance." Probably all would agree that it is important to forecast with reasonable accuracy the appearance of new Soviet weapons systems, and to do so well ahead of their initial operational dates. Probably most would agree further that the weapons systems mentioned in the following list were the most important to

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

forecast, though others might certainly be added. Here is the record of NIE's in this matter:

a. In 1950 (the first year of National Intelligence Estimates in present form), jet medium bombers were forecast for the Soviet forces in 1952; they appeared in 1954.

b. In 1951, thanks to the appearance of a single aircraft identified as a heavy bomber (the so-called Type 31, never thereafter seen) heavy bombers were thenceforth estimated to be brought into Soviet forces; they were in 1954.

c. In October 1953, an NIE said that a Soviet surface-to-air missile of native design could be developed by 1955; the first SA-1 missiles (based on a German design) became operational around Moscow in 1953; all sites were operational by 1956. The first SA-2 battalion became operational in 1958 or early 1959.

d. In October 1954, an NIE said that the Soviets could have an ICBM ready for series production about 1963, or at the earliest possible date in 1960; the SS-6 became operational in 1960.

e. In 1957, an NIE said that the Soviets could not have an ABM by 1962. In 1959 the estimate was that the Soviets were pushing hard on the problem and could have a first operational capability with an ABM in the period 1963-1966; the Moscow ABM system began to be operational in 1968.

f. In 1965, an NIE said that the Soviets would probably produce a new class of ballistic missile submarine, that it would almost certainly be nuclear powered, and that it would carry perhaps 6-12 missiles of an improved type. That NIE also judged that a new missile with about 1,000 n.m. range would come into service in 1967-1968. These estimates were made purely on the basis of Soviet requirements; there was no hard evidence of such developments at the time. In 1966 we saw the first unit of the new Y-class submarine having 16 launch tubes, and the Soviets began testing a new missile with an estimated range of 1,300 n.m.; this system—submarine and missile—became operational in 1968.

g. In 1965, an NIE said that the Soviets could probably attain an operational capability with a multiple independently guided re-entry vehicle (MIRV) in the period 1970-1975.

I think it true to say that in the past fifteen or twenty years no important new Soviet weapons system has appeared which had not been heralded in advance in National Intelligence Estimates. The initial operational dates have often been wrong, but as the above citations indicate they have usually been wrong because they have set the date too early; they have not "erred far more frequently on the conservative side than otherwise."

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9

To attack Senator Dodd's contentions is not to prove anything conclusively about the validity of National Intelligence Estimates as a whole. There are a good many people within the intelligence community (and probably outside as well) who feel that the net impact of NIE's over the years has been to over- rather than under-estimate Soviet military capabilities and intentions. If one of these persons were to draw up a documented indictment, it could probably be countered in the same fashion that I have tried to counter Senator Dodd's charges; and still nothing would be finally demonstrated. The estimator himself finds it useful to look into his record, not merely for the satisfaction or chagrin he may derive from the exercise, but because it may help him improve his performance in time to come. But the man whose opinion counts most—the "high-level policy-maker"—will never get his evaluation of NIE's from an exhaustive study of them. He will have no more than a vague impression—an impression, however, which will suddenly and emphatically crystallize whenever an estimate crucial to his immediate concern proves wrong. Once his view is thus formed it may take a long time to change.

Approved For Release 2005/12/14 : CIA-RDP79R00971A000400030001-9